

GenCore version 5.1.3
(c) 1993 - 2003 Compugen Ltd.

11	30	100..0	150	20	AYY24796
12	30	100..0	233	20	AYY05700
13	30	100..0	297	16	AAR9599
14	30	100..0	348	12	AAR10196
15	30	100..0	356	20	AYY24795
16	30	100..0	369	12	AAR12892
17	30	100..0	371	12	AAR10195
18	30	100..0	372	12	AAR10197..
19	30	100..0	372	12	AAR10200..
20	30	100..0	372	21	AYY84006
21	30	100..0	374	12	AAR10198
22	30	100..0	384	21	AYY84007
23	30	100..0	391	12	AAR10199
24	30	100..0	401	21	AYY84004
25	30	100..0	413	20	AYY25020
26	30	100..0	413	21	AYY84005
27	30	100..0	413	23	AAB00102
28	30	100..0	414	12	AAR10194
29	30	100..0	414	13	AAR20202
30	30	100..0	414	15	AAR63120
31	30	100..0	414	20	AYY24194
32	30	100..0	414	20	AYY01556
33	30	100..0	414	20	AAW94664
34	30	100..0	414	20	AWW94655
35	30	100..0	414	20	AAW86113
36	30	100..0	414	20	AAB86114
37	30	100..0	414	21	AYY90282
38	30	100..0	414	21	AAB01235
39	30	100..0	414	21	AAB01236
40	30	100..0	414	22	ARB14940
41	30	100..0	415	21	AYY99553..
42	30	100..0	415	21	AYY50870
43	30	100..0	440	6	APR50620
44	30	100..0	440	11	AAR06317..
45	30	100..0	440	11	AAR06317..

ALIGNMENT

טבנרט

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	6	20 AAV25009	Streptokinase deri
2	30	100.0	6	23 AB880001	Streptokinase deri
3	30	100.0	16	23 AB880005	Streptokinase deri
4	30	100.0	18	20 AAV25013	Streptokinase deri
5	30	100.0	20	20 AAV25014	Streptokinase deri
6	30	100.0	20	23 AB880006	Streptokinase deri
7	30	100.0	21	20 AAV25015	Streptokinase deri
8	30	100.0	21	23 AB880007	Streptokinase deri
9	30	100.0	71	20 AAV01550	Antigenic epitope
10	30	100.0	146	20 AAV01558	Peptide correspond

DE Streptokinase derived peptide#1.

PN XX Streptokinase; cell death; apoptosis; necrosis; nootropic; neuroprotective; antiparkinsonian; anticonvulsant; cytostatic; antiinflammatory; antiarthritic; antirheumatic; cardiant; antiatherosclerotic; vasotropic; immunosuppressive; anti-HIV; dermatological; antidiabetic; antiaemic; virucide; ophthalmological; antiulcer; antibacterial; antiparasitic; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; Huntington's disease; cerebellar degeneration; neoplastic disorder; cancer; inflammatory disorder; arthritis; inflammatory joint disorders; cardiovascular disease; heart failure; atherosclerosis; myocardial reperfusion injury; immune disease; autoimmune disease; acquired immunodeficiency syndrome; AIDS; rheumatoid arthritis; systemic lupus erythematosus; diabetes mellitus; pernicious anaemia; myelodegenerative diseases; viral diseases; macular degeneration; cataract; Crohn's disease; ulcerative colitis; pancreatic disease; prion disease; aging.

XX PA (RABK/) RABKIN S W.

XX PT Peptides that ameliorate cell death useful for treating conditions associated with cellular differentiation

XX PT

XX PS Claim 2; Column 12; 15pp; English.

XX CC AAV25009-Y25019 are novel peptides derived from streptokinase that ameliorate cell death. The products of the invention and their encoding nucleic acids may be useful for treating diseases and conditions related to aging, cellular differentiation, physical insult (e.g. physical trauma, anoxia, hyperthermia, hypothermia, chemical induced damage, and trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin and any other organ), viral disorders (e.g. hepatitis, retroviral infections, viral encephalitis, and AIDS/HIV), neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's disease, cerebellar degenerations, and familial amyotrophic lateral sclerosis (FALS)), cardiovascular disease (e.g. atherosclerosis, myocardial infarction, heart failure, cardiomyopathy, myocardial reperfusion injury, and hypertensive heart disease), immune disease (e.g. rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, erythema nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis, Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome, and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia, sarcomas, melanomas, carcinomas, neuroomas, cancers of the breast, brain, colon, cervix, and prostate, Hodgkin's disease and non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint disorders and inflammatory induced cell damage to the eye, brain and other organs), ischemia or reperfusion injury (e.g. myocardial ischemia and/or reperfusion injury, renal ischemia, spinal cord ischemia and/or reperfusion injury, retinal ischemia or infarction, and toxic damage to other organs from chemicals, radiation, and other noxious substances), macular degeneration, cataract formation, pancreatitis, Crohn's disease, ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor neuron diseases, degeneration of the spinal cord, Guillan Barré Syndrome and demyelinating disease), bypass surgery, chemotherapy, chemically-induced reperfusion, and therapeutics such as clozapine, AZT, and anthracyclines.

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 20; Length 6;
Best Local Similarity 100.0%; Pd. No. 7.8e+05; Indels 0; Gaps 0;

QY 1 SVDVEY 6
||| |
Db 1 SVDVEY 6

RESULT 2
ID ABB80001 standard; peptide; 6 AA.
XX AC ABB80001;
XX DR 26-JUL-2002 (first entry)

SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 23; Length 6;
Best Local Similarity 100.0%; Pd. No. 7.8e+05; Indels 0; Gaps 0;

nucleic acids may be useful for treating diseases and conditions related to aging, cellular differentiation, physical insult (e.g. physical trauma, anoxia, hyperthermia, hypothermia, chemically induced damage, and any other organ), viral disorders (e.g. hepatitis, heart, lungs, liver, skin infections, viral encephalitis, and AIDS/HIV), neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's disease, cerebellar degenerations, and familial amyotrophic lateral sclerosis (FALS)), cardiovascular disease (e.g. myocardial infarction, heart failure, cardiomyopathy, myocardial reperfusion injury, and hypertensive heart disease), immune disease (e.g. rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent diabetes mellitus, lupus, pernicious anemia, dermatomyositis, erythema nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis, Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome, and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia, sarcomas, myelomas, carcinomas, melanoma, cancers of the breast, brain, colon, cervix, and prostate, Hodgkin's disease and non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint disorders and inflammatory induced cell damage to the eye, brain and other organs), ischemia or reperfusion injury (e.g. myocardial ischemia and reperfusion injury, renal ischemia, spinal cord ischemia and/or reperfusion injury, retinal ischemia or infarction, and stroke), toxic insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other organs from chemicals, radiation, and other noxious substances), macular degeneration, cataract formation, pancreatitis, Crohn's disease, ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor neuron diseases, degeneration of the spinal cord, Guillain Bare Syndrome and demyelinating disease), bypass surgery, chemotherapy, clozapine, AZT, chemically-induced reperfusion, and therapeutics such as clozapine, AZT, CC and anthracyclines.

XX Sequence 18 AA;

Query Match Score 30; DB 20; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 SVDVY 6
  |||||
Db 1 SVDVY 6
```

RESULT 5
ID AAY25014 Standard; peptide: 20 AA.
AC
XX 23-AUG-1999 (first entry)

DB Streptokinase derived peptide 6 for amelioration of cell death.

KW Streptokinase; cell death; amelioration; treatment; disease; aging; AIDS; cellular differentiation; sclerosis; atherosclerosis; heart disease; heart failure; heart disease; cardiomopathy; hyperthermia; myocardial; reperfusion injury; immune disease; spinal cord; kidney; heart; lung; liver; skin; HIV; viral disorder; hepatitis; retroviral; infection; encephalitis; FALS; neurodegenerative disorder; Parkinson's disease; Alzheimer's disease; Huntington's disease; cerebellar degeneration; cardiovascular disease; familial amyotrophic lateral sclerosis; atherosclerosis; heart disease; heart failure; infarction; heart disease; cardiomopathy; hyperthermia; myocardial; reperfusion injury; immune disease; spinal cord; kidney; heart; lung; liver; skin; HIV; systemic lupus erythematosus; insulin-dependent diabetes mellitus; pernicious anemia; dermatomyositis; erythema nodosum; myasthenia gravis; Sjogren's syndrome; temporal arthritis; autoimmune polyarthritides; Wegener's granulomatosis; glomerulonephritis; anti-phospholipid syndrome; neoplastic disorder; leukemia; sarcoma; myeloma; carcinoma; neurooma; melanoma; cancer; breast; colon; cervix; prostate; stroke; Hodgkin's disease; non-Hodgkin's lymphoma; inflammatory disorder; stroke; ischemia; spinal cord; toxic insult; pulmonary; macular degeneration; cataract; pancreatic disease; Crohn's disease; ulcerative colitis; motor neuron disease; Guillain Bare Syndrome; demyelinating disease; bypass surgery; chemotherapy; clozapine; AZT; anthracycline.

XX OS Synthetic.
XX PN US5917013 A.
XX PD 29-JUN-1999.
XX PF 05-DEC-1996; 96US-0759599.
XX PR 06-DEC-1995; 95US-0003233.
XX PR 05-DEC-1996; 96US-0759599.
PA (RABK/) RABKIN S W.
PA Krystral G, Rabkin SW;
PA XX
PS XX
DR WPI; 1999-394231/33.

Claim 7, Column 12; 15pp; English.
XX Peptides that ameliorate cell death useful for treating conditions associated with cellular differentiation
XX
Claim 7, Column 12; 15pp; English.
XX
AAY25009-Y25019 are novel peptides derived from streptokinase that ameliorate cell death. The products of the invention and their encoding nucleic acids may be useful for treating diseases and conditions related to aging, cellular differentiation, physical insult (e.g. physical trauma, anoxia, hyperthermia, hypothermia, chemically induced damage, and trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin and any other organ), viral disorders (e.g. hepatitis, retroviral, infections, viral encephalitis, and AIDS/HIV), neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's disease, cerebellar degenerations, and familial amyotrophic lateral sclerosis (FALS)), cardiovascular disease (e.g. atherosclerosis, myocardial infarction, heart failure, cardiomyopathy, myocardial reperfusion injury, and hypertensive heart disease), immune disease (e.g. diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, myasthenia gravis, Sjogren's syndrome, temporal arthritis, myasthenia gravis, Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome, and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia, sarcomas, myelomas, carcinomas, neuroomas, melanoma, cancers of the breast, brain, colon, cervix, and prostate, Hodgkin's disease and non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint disorders and inflammatory induced cell damage to the eye, brain and other organs), ischemia or reperfusion injury (e.g. myocardial ischemia and reperfusion injury, renal ischemia, spinal cord ischemia and/or reperfusion injury, retinal ischemia or infarction, and stroke), toxic insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other organs from chemicals, radiation, and other noxious substances), macular degeneration, cataract formation, pancreatitis, Crohn's disease, ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor neuron disease), degeneration of the spinal cord, Guillain Bare Syndrome and demyelinating disease), bypass surgery, chemotherapy, clozapine, AZT, CC and anthracyclines.

XX Sequence 20 AA;

Query Match Score 30; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 SVDVY 6
  |||||
Db 1 SVDVY 6
```

RESULT 6
ID ABB80006 standard; peptide: 20 AA.
AC
XX

AC ABB80006;

ID ABB80006 standard; peptide: 20 AA.
XX

AC ABB80006;

reperfusion injury, and hypertensive heart disease), immune disease (e.g. rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent diabetes mellitus, lupus, pernicious anemia, dermatomyositis, erythema nodosum, Sjögren's syndrome, temporal arthritis, myasthenia gravis, Wegener's Granulomatosis, glomerulonephritis, anti-phospholipid syndrome, and autoimmune polyarthritis), a neoplastic disorder (e.g. leukemia, sarcomas, melanomas, carcinomas, neuroomas, cancers of the breast, brain, colon, cervix, and prostate, Hodgkin's disease and non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint disorders and inflammatory induced cell damage to the eye, brain and other organs), ischemia or reperfusion injury (e.g. myocardial ischemia and/or reperfusion injury, renal ischemia, spinal cord ischemia and/or reperfusion injury, retinal ischemia or infarction, and stroke), toxic insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other organs from chemicals, radiation, and other noxious substances), macular degeneration, cataract formation, pancreatitis, Crohn's disease, ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor neuron diseases, degeneration of the spinal cord, Guillain Barré syndrome and demyelinating disease), bypass surgery, chemotherapy, and chemically-induced reperfusion, and therapeutics such as clozapine, AZT, and anthracyclines.

Sequence	21 AA:	Query Match	100.0%	Score	30;	DB	20;	Length	21;
		Best Local Similarity	100.0%	Pred.	No.	2.1;			
		Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;
								Gaps	0;
1	SYDVEY	6							
2	SYDVEY	6							

RESULT 8
BBBBB0007 command: norm:do 21 23

ABB80007;

8-005-2002 (תינוקות)

treptokinase; cell death; apoptosis; necrosis; nootropic;
neuroprotective; anticarsonian; anticonvulsant; cytoclastic;
antiinflammatory; antiarthritic; antihypertensive; cardiotonic;
antiatherosclerotic; vasotropic; immunosuppressive; anti-HIV;
antidiabetic; antiaemic; virucide; ophthalmological;
antimicrobial; antibacterial; antiparasitic; neurodegenerative disease;
antibacterial; Alzheimer's disease; Huntington's disease;
inflammatory disorder; neoplastic disorder; cancer;
inflammatory disorder; arthritis; inflammatory joint disorders;
cardiovascular disease; heart failure; atherosclerosis;
myocardial reperfusion injury; immune disease; autoimmune disease;
acquired immunodeficiency syndrome; AIDS; rheumatoid arthritis;
systemic lupus erythematosus; diabetes mellitus; pernicious anaemia;
telangiectatic disease; viral diseases; macular degeneration;
ataxiat; Crohn's disease; ulcerative colitis; ulcerative colitis;
pancreatitis; aging.

synthetic.

SS6348567-B1.

99-APR-1999; 99US-0294457.

6-DEC-1995; 95US-008233P.
5-DEC-1996; 96US-0759599.

(MOLE-) MOLECULAR THERAPEUTICS INC.

PI Krystal G, Rabkin SW;
XX WPI; 2002-266542/31.
DR

PT New peptides obtained from streptokinase, useful in ameliorating cell death due to apoptosis and/or necrosis and treating neurodegenerative, neoplastic, immune, cardiovascular and inflammatory disorders -
XX
PS Claim 9; Column 5; 18PDP; English.

The invention relates to an isolated peptide obtained from streptokinase.

or its derivative or analog, which ameliorate cell death. The activity of peptides of the invention may be described as, nootropic, neuroprotective, antiparkinsonian, anticonvulsant, cytostatic, antilinflammatory, antiarthritic, antiallergic, antianemic, virucide, antihistaminic, antidiabetic, antianabolic, immunosuppressive, anti-HIV, antithrombotic, antibacterial and antiparasitic. Peptides of the invention ameliorates apoptosis and necrosis in warm-blooded animal. Compositions comprising peptides of the invention are useful for treating neurodegenerative diseases (e.g. Parkinson's, Alzheimer's, Huntington's disease and cerebellar degeneration) neoplastic disorders including cancer, inflammatory disorders (e.g. arthritis, inflammatory joint disorders), cardiovascular diseases (e.g. heart failure, atherosclerosis and myocardial reperfusion injury), immune diseases (e.g. autoimmune disease, acquired immunodeficiency syndrome (AIDS), rheumatoid arthritis, systemic lupus erythematosus, diabetes mellitus, pernicious anaemia), myeloidnegenerative diseases, viral diseases, and degenerative diseases of any organ. Other disorders include macular degeneration, cataracts, Crohn's disease, ulcerative colitis, parasites, pancreatic diseases, infectious diseases including bacteria, parasite, prion-based diseases, and accelerated aging. The current sequence represents a streptokinase derived peptide of the invention with an ability to ameliorate cell death in cardiac myocytes.

SO Sequence 21 AA:

Query	Match	Score 30:	DB 23:	Length 21:
Best Local Similarity	100.0%	Score 30: 100.0%	DB 23: Pred. No. 2.1.	Length 21: 0
Best Matches	100.0%	Score 30: 100.0%	DB 23: Mismatches 0;	Length 21: 0
6; Conservative	0;	Score 30: 0;	DB 23: Indels 0;	Length 21: 0
		Score 30: 0;	DB 23: Gaps 0;	Length 21: 0

RESULT 9
AYA01550 standard; peptide; 71 AA.
ID AYA01550
XX

XX	XX	XX	XX	XX
XX	DT	18-JUN-1999	(first entry)	
XX	XX	XX	Antigenic epitope of streptokinase, spanning amino acids 138-208.	
XX	XX	XX	Antigenic Peptide; streptokinase; streptokinase-specific antibody	
XX	XX	XX	thrombolytic activity; thrombolytic therapy; glomerulonephritis;	
XX	XX	XX	rheumatic fever.	

Streptococcus equisimilis.
OS XX PN WO9908698-A1.
XX

XX 18-AUG-1998; 98WO-US17114
 PP XX 18-AUG-1997; 9705-0055911
 PR XX (GEN HOSPITAL CORP.
 PA. (HARD) HARVARD COLLEGE.
 PA. XX

PI Parhami-Seren B, Reed GL;
 XX WPI; 1999-190113/16.
 XX New polypeptides which bind streptokinase-specific antibodies -
 PT useful in thrombolytic therapy
 XX
 PT
 XX
 PS Claim 16; Page 5; 44pp; English.

Sequence 71 AA;

Query Match 100.0%; Score 30; DB 20; Length 71;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 Db 20 SVDVEY 25

Query Match 100.0%; Score 30; DB 20; Length 71;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 Db 10 SVDVEY 15

Query Match 100.0%; Score 30; DB 20; Length 146 AA;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 Db 10 SVDVEY 15

RESULT 11
 AAY24796
 ID AAY24796 standard; Protein; 150 AA.
 AC AAY24796;
 DT 26-AUG-1999 (first entry)
 XX Recombinant streptokinase rSKI44-293.
 DB Streptococcus equisimilis.
 XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nsK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX OS Streptococcus equisimilis.
 XX Synthetic.
 XX PN WO9931247-A1.
 XX PD 24-JUN-1999.
 XX XX
 XX PF 15-DEC-1998; 98WO-US26694.
 XX PR 15-DEC-1997; 97US-0069497.
 XX PA (HARD) HARVARD COLLEGE.
 XX PI Reed GL;
 XX DR WPI; 1999-395183/33.
 XX DR N-PSDB; AAX80494.
 XX PT N-terminally deleted streptokinase
 XX
 PS Claim 34; Page 67-68; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein;
 CC a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising
 CC a bacterial fibrin-dependent plasminogen activator is useful for
 CC dissolving blood clots in patients with a thrombotic condition, e.g.
 CC myocardial infarction, venous thrombosis, pulmonary embolism, cerebral
 CC thrombosis, graft thrombosis and arterial thrombosis. The modified
 CC streptokinase can also be used in non-human mammals. Streptokinase
 CC activation of plasminogen is at least 10-fold, preferably 100-fold
 CC greater in the presence of fibrin than in the absence of fibrin. The
 CC modified streptokinase has at least one amino acid substitution that
 CC inactivates substrate site for proteolytic cleavage. This reduces the
 CC rate of degradation of the streptokinase at least two-fold. The present
 CC sequence represents recombinant streptokinase, designated rSKI44-293.
 XX Sequence 150 AA;

PI Parhami-Seren B, Reed GL;
 XX WPI; 1999-190113/16.
 XX New polypeptides which bind streptokinase-specific antibodies -
 PT useful in thrombolytic therapy
 XX
 PS Disclosure; Page 13; 44pp; English.
 XX The present sequence represents a peptide corresponding to
 CC amino acids 148-293 of streptokinase. The specification describes a
 CC polypeptide which binds to a streptokinase-specific antibody and
 CC prevents the antibody binding to native streptokinase. The specification
 CC also describes a synthetic polypeptide (P1) comprising an epitope which
 CC binds to an streptokinase-specific antibody and reduces thrombolytic
 CC activity of streptokinase. P1 is used in thrombolytic therapy, and to
 CC prevent or treat glomerulonephritis and rheumatic fever.

Query Match 100.0%; Score 30; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Streptokinase from pST-1.
 XX
 KW streptokinase; recombinant production; haemolytic Streptococcus.
 OS Synthetic.
 XX
 PN WO9527050-A1.
 XX
 PD 12-OCT-1995.
 XX
RESULT 12
 AAY01557 standard; peptide; 233 AA.
 ID AAY01557
 XX
 AC AAY01557;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Peptide corresponding to amino acids 120-352 of streptokinase.
 XX
 KW Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KW rheumatic fever.
 XX
 OS Streptococcus equisimilis.
 XX
 PN WO9308698-A1.
 XX
 PD 25-FEB-1999.
 XX
 PP 18-AUG-1998; 98WO-US17114.
 XX
 PR 18-AUG-1997; 97US-0055911.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Parhami-Seren B, Reed GL;
 XX
 DR 1999-190113/16.
 XX
 PS New polypeptides which bind streptokinase-specific antibodies -
 PT useful in thrombolytic therapy
 XX
 Disclosure: Page 13; 44PP; English.

XX
 CC The present sequence represents a peptide corresponding to
 CC amino acids 120-352 of streptokinase. The specification describes a
 CC polypeptide which binds to a streptokinase-specific antibody and
 CC prevents the antibody binding to native streptokinase. The specification
 CC also describes a synthetic polypeptide (P1) comprising an epitope which
 CC binds to an streptokinase-specific antibody and reduces thrombolytic
 CC activity of streptokinase. P1 is used in thrombolytic therapy, and to
 CC prevent or treat glomerulonephritis and rheumatic fever.
 XX
 SQ Sequence 233 AA;

Query Match 100.0%; Score 30; DB 20; Length 233;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Streptokinase from pST-1.
 XX
 KW streptokinase; recombinant production; haemolytic Streptococcus.
 OS Synthetic.
 XX
 PN (UYSH-) UNIV SHANGHAI MEDICAL.
 XX
 PT Prep'n. of streptokinase - useful for treating thrombus diseases.
 XX
 PS Example 2; Page 6-8; 24pp; Chinese.
 XX
RESULT 13
 AAR1599 standard; protein; 297 AA.
 ID AAR1599
 XX
 AC AAR1599;
 XX
 DT 08-AUG-1996 (first entry)
 XX
 PR 11-APR-1990; 90JP-0096830.
 PR 11-JUL-1989; 89JP-0179432.
 PR 27-NOV-1989; 89JP-0307957.
 XX
 PA (SAKA) OTSUKA PHARM FACTOR.
 XX
 PI Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
 XX
 DR WPI; 1991-016179/03.

XX Synthetic gene encoding streptokinase - scale, high purity prodn.
PT of streptokinase used as a thrombolytic agent
XX Claim 7; Page 59; 76pp; English.

XX Streptokinase derivative having a C-terminal deletion (i.e. Pro 373-Lys 414) and Arg 45 to Gly 68 also being deleted. Derivatives show a decreased antigenicity, higher stability in blood and greater selectivity of thrombolytic activity and specificity. The protein is used as a thrombolytic agent in patients with lung thrombus or myocardial infarction. See also AAQ10230, AARI0195 and AARI0197-R10200.

XX Sequence 348 AA;
SQ Score 100.0%; DB 12; Length 348;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
Qy 1 SVDVY 6
Db 133 SVDVY 138

XX Recombinant streptokinase rSK59-414.
XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator; NSK; rSK; bacterial; blood clot; thrombotic condition; myocardial infarction; venous thrombosis; pulmonary embolism; cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX Streptococcus equisimilis.
OS Synthetic.
XX PN WO9931247-A1.
XX PD 24-JUN-1999.
XX PF 15-DEC-1998; 98WO-US26694.
XX PR 15-DEC-1997; 97US-0069497.
XX (HARD) HARVARD COLLEGE.
PI Reed GL;
XX WPI; 1999-395183/33.
DR N-PSDB; AAX80493.

XX N-terminally deleted streptokinase
XX Claim 35; Page 65-66; 73pp; English.
CC The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also isolated are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for

XX dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold. The greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence represents recombinant streptokinase, designated rSK59-414.

XX SQ Sequence 356 AA;
Query Match 100.0%; Score 30; DB 20; Length 356;
Best Local Similarity 100.0%; Prod. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Qy 1 SVDVY 6
Db 99 SVDVY 104

Search completed: January 10, 2003, 11:46:51
Job time : 36 secs

RESULT 15
AYZ4795
ID AYZ4795 standard; Protein: 356 AA.
XX AC AYZ4795;
XX DT 26-AUG-1999 (First entry)

XX Recombinant streptokinase rSK59-414.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator; NSK; rSK; bacterial; blood clot; thrombotic condition; myocardial infarction; venous thrombosis; pulmonary embolism; cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX Streptococcus equisimilis.
OS Synthetic.
XX PN WO9931247-A1.
XX PD 24-JUN-1999.
XX PF 15-DEC-1998; 98WO-US26694.
XX PR 15-DEC-1997; 97US-0069497.
XX (HARD) HARVARD COLLEGE.
PI Reed GL;
XX WPI; 1999-395183/33.
DR N-PSDB; AAX80493.

XX N-terminally deleted streptokinase

XX PS Claim 35; Page 65-66; 73pp; English.
CC The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also isolated are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for

escherichia
 P75800 escherichia
 aeropyrum p
 P12435 bluetongue
 Q65749 bluetongue
 Q65750 bluetongue
 P03539 bluetongue
 P20608 bluetongue
 P56582 bluetongue
 P56748 bluetongue
 P25167 drosophila
 P47170 saccharomyces
 P70107 carica
 Q10470 cayenne
 US-09-919-703-1
 title:
 run on: January 10, 2003, 11:44:11 ; Search time 10 Seconds
 (without alignments)
 24.886 Million cell updates/sec

sequence: 1 SVDVRY 6

ALIGNMENTS					
scoring table:	BLOSUM62	searched:	Gapop 10.0 , Gapext 0.5	total number of hits satisfying chosen parameters:	112892
minimum DB seq length:	0	maximum DB seq length:	200000000	post-processing:	Minimum Match 0% Maximum Match 100%
					Listing first 45 summaries
					SwissProt_40-*
RESULT 1					
LBC2	MEDTR	ID	LEC2	MEDTR	STANDARD:
AC	Q01807;	DT	01-JUL-1993	(Rel. 26, Created)	PRT: 280 AA.
DT	01-JUL-1993	DT	01-JUL-1993	(Rel. 26, Last sequence update)	
DT	15-JUL-1999	DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DB		DB		Truncated lectin 2 precursor.	
GN	LEC2				
OS	Medicago truncatula (Barrel medic).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots; Rosids				
OC	euroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicaceae				

SUMMARIES						
result No.	Query Score	Match Length	DB ID	Description		
1	30	100.0	280	1 LEC2 MEDTR	001807 medicago tr	RN
2	30	100.0	440	1 STRP_STREO	P00779 streptococc	RP
3	30	100.0	440	1 STRP_STREY	P10519 streptococc	RC
4	30	100.0	440	1 STRP_STRSP	P81307 methanococc	RA
5	29	96.7	183	1 Y3AA_MENIA	P06567 bacillus su	RT
6	28	93.3	311	1 DNAL_BACSU	089023 mus musculu	RL
7	27	90.0	562	1 TPP1_MOUSE	Q9xb88 canis famili	CC
8	27	90.0	563	1 TPP1_CANFA	Q9eqv6 rattus norv	CC
9	27	90.0	563	1 TPP1_RAT	P03764 bacteriophag	CC
10	27	90.0	774	1 STP_LAMBD	Q92442 mucor javan	CC
11	27	90.0	864	1 AGIIJ_MUCIJA	P76072 escherichia	CC
12	27	90.0	1120	1 TIPR_ECOLI	P32492 saccharomyces	CC
13	27	90.0	1471	1 MYS4 YEAST	001901 p genome po	CC
14	27	90.0	3344	1 P0LG_PRSVH	P78265 escherichia	CC
15	26	86.7	47	1 YORP_ECOLI	P41161 haemophilus	DR
16	26	86.7	148	1 YCBG_HABIN	Q9spqr2 dictyosteli	DR
17	26	86.7	205	1 RAC1_DICDI	Q91zts arabidopsis	DR
18	26	86.7	263	1 EXL3_ARATH	Q9yqj0 aeropyrum	DR
19	26	86.7	305	1 HEN3_AEROB	Q49425 mycoplasma	DR
20	26	86.7	307	1 RUVB_MYCOB	P73242 mycoplasma	DR
21	26	86.7	307	1 RUVB_MYCOP	Q44457 agrobacteri	DR
22	26	86.7	321	1 FLIM_AGENT5	P22825 myxoma viru	DR
23	26	86.7	326	1 VT2_MYXV	Q59498 lactococcus	DR
24	26	86.7	341	1 RBCA_LACLA	P02680 ractus norv	KW
25	26	86.7	445	1 FIBG_RAT	Q65332 mycobacteri	FT
26	26	86.7	463	1 SYG_MYCTU	P997b3 thermotoga	FT
27	26	86.7	496	1 ARAB_THEMEA	Q84786 chlamydia t	FT
28	26	86.7	526	1 SYK_CHLTR	O84898 mycoplasma	FT
29	26	86.7	571	1 UVRG_MYCBV	Q10115 haemophilus	FT
30	26	86.7	611	1 ACE_HAEIE	Q10114 drosophila	FT
31	26	86.7	615	1 ACE_DROME	Q96617 schizosacch	FT
32	26	86.7	693	1 RAPI_SCHIZO	PS2303 bacillus en	FT
33	26	86.7	749	1 YPRA_BACTIV	PS2303 bacillus en	FT

Query Match Score 30; DB 1; Length 280;
 Best Local Similarity 100.0%; Pred. No. 7.5; O; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

STRP_STREQ STANDARD; PRT; 440 AA.

Y 1 SVDVEY 6
 |||||
 208 SVDVEY 213

STREPTOKINASE C precursor.

NCBI_TaxID=119602;
 [1]

IP SEQUENCE FROM N.A.
 STRAIN=HA6A;
 MEDLINE=85232082; PubMed=2989113;
 MA Like H., Roe B., Ferretti J.J.;
 "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis HA6A";
 Gene 34:357-362(1985).
 [2]

IP SEQUENCE OF 27-440.
 MEDLINE=93127125; PubMed=6760891;
 JACKSON K.W., TANG J.J.;
 "Complete amino acid sequence of streptokinase and its homology with serine proteases.",
 Biochemistry 21:6620-6625(1982).

-I- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR, IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE INVASIVENESS OF THE CELLS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; K02986; AAA2674.1; -.
 EMBL; X72832; CAA51351.1; -.
 PIR; A00587; B2S0.
 PIR; A22801; A22801.
 InterPro; IPR004093; Staphylokinase.
 Pfam; PF0284; Staphylokinase_3.
 Plasminogen activation; Signal; Virulence.

T SIGNAL 1 26 STREPTOKINASE C.
 T CHAIN 27 440 STREPTOKINASE A.
 T VARIANT 195 195 L > D.
 T VARIANT 207 207 D > L.
 T CONFLICT 298 300 EKY > LEYK (IN REF. 2).
 T CONFLICT 438 438 N > D (IN REF. 2).
 Q SEQUENCE 440 AA; 50140 MW; 8FC1F22648ACCC7A CRC64;

Query Match Score 30; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 12; O; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 |||||
 183 SVDVEY 188

RESULT 3

STRP_STRPY STANDARD; PRT; 440 AA.

AC P10520;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Streptokinase A precursor.
 GN SKA OR SKY1979.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP130713 / Serotype M1;
 RX MEDLINE=9116264; PubMed=2616590;
 RA Walter F., Siegel M., Malke H.;
 RA "Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes type 1 strain."
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 RL Nucleic Acids Res. 17:1261-1261(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192688; PubMed=11295296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Yuan Y., Jia H.G., Najar P.Z., Ren Q., White J., Clifton S.W., Roe B.A., McLaughlin R., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 -I- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR, IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE INVASIVENESS OF THE CELLS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X13399; CAA1765.1; -.
 DR Pfam; PF02821; Staphylokinase_3.
 DR Plasminogen activation; Signal; Virulence; Complete proteome.
 DR SIGNAL 1 26 STREPTOKINASE A.
 FT CHAIN 27 440 STREPTOKINASE A.
 FT CONFLICT 163 L > V (IN REF. 1).
 FT CONFLICT 345 R > G (IN REF. 1).
 FT CONFLICT 373 D > N (IN REF. 1).
 FT CONFLICT 428 D > Y (IN REF. 1).
 FT CONFLICT 438 438 N > D (IN REF. 1).
 SQ SEQUENCE 440 AA; 49924 MW; D62278E040B758DB_CRC64;

Query Match Score 30; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 12; O; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 |||||

Db	183	SVDNEY 188	
RESULT 4			
STRP STRSP	STANDARD;	PRT;	440 AA.
P10519; AC			
DT 01-JUN-1989 (Rel. 11, Created)			
DT 01-JUN-1989 (Rel. 11, Last sequence update)			
DT 01-NOV-1997 (Rel. 35, Last annotation update)			
DB Streptokinase G precursor.			
GN SKG			
OS Streptococcus sp. (strain 19909).			
CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC Streptococcus.			
OX NCBI_TaxID=1306;			
RN			
RP SEQUENCE FROM N.A.			
RX MEDLINE=89160265; PubMed=2922269;			
RA Walter F. Siegel M., Malke H.;			
RT "Nucleotide sequence of the streptokinase gene from a group-G streptococcus.";			
RT Nucleic Acids Res. 17:1262-1262 (1989).			
RL FUNCTION. THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR, IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE INVASIVENESS OF THE CELLS.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC EMBL: X13400; CAA31766 1; -.			
DR PIR: S02723; S02723;			
DR InterPro: IPR004053; staphylokinase.			
DR Pfam: PF02821; Staphylokinase_3.			
KW Plasminogen activation; Signal; Virulence.			
FT SIGNAL	1	26	
FT CHAIN	27	440	
SEQUENCE	440 AA;	50199 MW;	5521F8825FEB16EA CRC64 ;
Query Match Score 30%; DB 1; Length 440;			
Best Local Similarity 100.0%; Pred. No. 12;			
Matches 6; Conservative 0; Mismatches 0; Gaps 0;			
RN			
Qy 1 SVDNEY 6			
Db 183 SVDNEY 188			
RESULT 5			
Y34A_METJJA	STANDARD;	PRT;	183 AA.
ID Y34A_METJJA			
AC P81307;			
DT 15-JUL-1998 (Rel. 36, Created)			
DT 15-JUL-1998 (Rel. 36, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Hypothetical protein. Mj0347.1.			
GN Methanococcus jannaschii.			
OS Archaea; Euryarchaeota; Methanococci; Methanococcales;			
OC Methanococcaceae; Methanococcaceae.			
OX NCBI_TaxID=2190;			
RN			
RP SEQUENCE FROM N.A.			
RC Methane-16S; PubMed=95387221;			
RX MEDLINE=98048457;			
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,			
RA Sanders J., Zimmerman P.T., Harwood C.R.;			
RT "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus subtilis chromosome containing genes responsible for stress responses, the utilization of plant cell walls and primary metabolism.";			
RT Genes on the Bacillus subtilis chromosome. 14:999-9999(1986).;			
RL Microbiology 142:3067-3078(1996).			
RN [3]			
RP SEQUENCE FROM N.A.			
RC STRAIN=16S;			
RX MEDLINE=97124191; PubMed=8969504;			
RA Ogasawara N., Moriya S., Mazza P.G., Yoshikawa H.;			
RT "Nucleotide sequence and organization of dnaB gene and neighbouring genes on the Bacillus subtilis chromosome.";			
RT Genes on the Bacillus subtilis chromosome. 14:999-9999(1986).;			
RL Nucleic Acids Res. 14:999-9999(1986).;			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=16S;			
RX MEDLINE=97124191; PubMed=8969504;			
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,			
RA Sanders J., Zimmerman P.T., Harwood C.R.;			
RT "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus subtilis chromosome containing genes responsible for stress responses, the utilization of plant cell walls and primary metabolism.";			
RT Genes on the Bacillus subtilis chromosome. 14:999-9999(1986).;			
RL Microbiology 142:3067-3078(1996).			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=16S;			
RX MEDLINE=98048457;			
RA Lapidus A., Galperin N., Sorek R., Ehrlich S.D.;			
RT "Sequencing and functional annotation of the Bacillus subtilis genome in the 200 kb rrnB-dnaB region.";			
RT Microbiology 143:3431-3441(1997).			
RN [4]			
RP SEQUENCE FROM N.A.			
RC STRAIN=16S;			
RX MEDLINE=98048457;			
RA Built C.J., White O., Olsen G.J., Zhou L., Pleischmann R.D.;			
RA Built C.J., Olsen G.J., Zhou L., Pleischmann R.D.;			

Sequence Similarity Search Results							
FT	CONFLICT	19	19	K -> N (IN REF. 5).			
RX	CONFLICT	24	24	M -> T (IN REF. 5).			
RA	SEQUENCE	311 AA;	36114 MW;	A86FC24AB641264 CRC64;			
RA	Query Match	93.3%	Score 28;	DB 1;	Length 311;		
RA	Best Local Similarity	66.7%	Pred. No. 25;				
RA	Matches	4;	Conservative	2;	Mismatches	0;	
RA	Qy	1 SYDVEY 6			Indels	0;	Gaps 0;
RA		1 : :					
RA	Db	93 SIDNEY 98					
RESULT 7							
TPPI_MOUSE	ID_TPP1_MOUSE	STANDARD;	PRT;	562 AA.			
AC	O9Q023; Q9QUS7;						
DT	15-DEC-1998 (Ref. 37, Created)						
DT	16-OCT-2001 (Ref. 40, Last sequence update)						
DT	15-JUN-2002 (Ref. 41, Last annotation update)						
DB	Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-I) (Tripeptidyl aminopeptidase) (Lysosomal peptidase insensitive protease) (LPIC).						
DB	CLN2 OR TPP1.						
GS	Mus musculus (Mouse).						
OC	Bukaryota; Metazoa; Chordata; Craniata; Buteleostomi;						
OC	Macromia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
NCBI_TaxID	10090;						
RN							
RP	SEQUENCE FROM N.A.						
RX	Medline:9912656; PubMed:9989590;						
RX	Vines D.J.; Warburton M.J.;						
RA	"Classical late infantile neuronal ceroid lipofuscinosis fibroblasts are deficient in lysosomal tripeptidyl peptidase I."						
RA	FEBS Lett. 443:131-135(1999).						
RN							
[2]	SEQUENCE FROM N.A.						
RX	Medline:20028316; PubMed:10556422;						
RA	Sleat D.E.; Lobel P.;						
RA	"Marine homologue of the lysosomal pepstatin insensitive protease which is deficient in human classical late infantile neuronal ceroid lipofuscinosis."						
RA	Submitted (DEC-1998) to the EMBL/GenBank/DBJ/DBU databases.						
RN							
[3]	SEQUENCE FROM N.A.						
RP	Medline:21085660; PubMed:11217851;						
RX	Arakawa T., Hara A., Shibusawa K., Yonohiro M., Itoh M., Ishii Y.,						
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,						
RA	Katz M.L., Liu P.-C., Grob-Nunn S.E., Shibuya H., Johnson G.S.;						
RA	"Characterization and chromosomal mapping of a mouse ortholog of the late-onset infantile ceroid-lipofuscinosis gene CLN2.,"						
RA	Mamm. Genome 10:1050-1053 (1999).						
RN							
[4]	SEQUENCE FROM N.A.						
RP	Medline:92195821; PubMed:1549481;						
RA	Koonin E.V.;						
RA	"Archaeabacterial virus SSV1 encodes a putative DnaA-like protein.";						
RA	Proc. Natl. Acad. Sci. U.S.A. 84:653-657(1987).						
RN							
RP	SIMILARITY TO DNA.						
RX	Medline:95291463; PubMed:7773414;						
RA	Bruand C., Ehrlrich S.D.;						
RA	"The Bacillus subtilis dnaB gene is part of the dnaB operon.,"						
RA	Microbiology 141:1199-1200(1995).						
CC	-- SIMILARITY: SOME, TO DNA FROM VARIOUS BACTERIA.						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).						
RN							
RP	IDENTIFICATION						
RX	Medline:92195821; PubMed:1549481;						
RA	"The Bacillus subtilis dnaB gene is part of the dnaB operon.,"						
RA	Microbiology 141:1199-1200(1995).						
CC	-- SIMILARITY: PROBABLY INVOLVED IN DNA REPLICATION.						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).						
CC	EMBL: X04963; CAA2853.1; -.						
CC	EMBL: AF008220; AAC0359.1; -.						
DR	EMBL: Z99118; CAB14838.1; -.						
DR	EMBL: M15183; AAA22405.1; -.						
DR	PIR: B24720; IOBS44.						
DR	PIR: C26580; C26580.						
DR	Primesome, DNAreplication, ATP-binding; Complete proteome.						
KW	Primesome, DNAreplication, ATP (PROBABLE).						
FT	NP-BIND	168	175	ATP	(PROBABLE).		
CC	FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I.						
CC	"Functional annotation of a full-length mouse cDNA collection."						
CC	Nature 409:685-690(2001)						
-1	FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I.						

CC activity. May act as a non-specific lysosomal peptidase which generates tripeptides from the breakdown products produced by lysosomal proteinases. Requires substrates with an unsubstituted N-terminus (By similarity).
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide, but also endopeptidase activity.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Activated by autocatalytic proteolytical processing upon acidification (By similarity).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: A0011912; CA09863_1; ALT_INIT.
 DR EMBL: AP124599; ADD32573_1; -.

DR EMBL: AP111172; APD03003_1; -.

DR MEROPS; S53_003; -.

DR MGDB: I1336194; Cln2.

KW Hydrolase; Protease; Serine protease; Zymogen; Signal; Lysosome;

KW Glycoprotein.

FT SIGNAL 1 19 BY SIMILARITY.

FT PROPEP 20 19 REMOVED IN MATURE FORM (BY SIMILARITY).

FT CHAIN 195 562 TRIPETIDYL-PePTIDAS E I.

FT ACT_SITE 359 359 BY SIMILARITY.

FT ACT_SITE 474 474 NUCLEOPHILE (BY SIMILARITY).

FT ACT_SITE 516 516 BY SIMILARITY.

FT CARBOHYD 209 209 BY SIMILARITY.

FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 1 1 M-> V (IN REF. 1).

FT CONFLICT 562 562 P-> LDPPVPP (IN REF. 1).

SQ SEQUENCE 562 AA: 61342 MW; OAF8167EA1A63396 CRC64;

Query Match 90.0% Score 27; DB 1; Length 563;

Best Local Similarity 83.3%; Pred. No. 83; 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 9 TPP1_RAT STANDARD; PRT; 563 AA.

ID TPP1_RAT Q9E0V6; AC DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DB Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-I) (Tripeptidyl aminopeptidase) (Lysosomal peptidin insensitive protease) (LPIC).

GN CLN2 OR TPP1.

OS Rattus norvegicus (Rat).

OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi; NCBI_TaxID=10116;

RN 11] RSEQUENCE FROM N.A.

RC TISSUE-Liver;

RA Du P., Kato S., Li Y., Maeda T., Yamane T., Yamamoto S., Fujiwara M., Yamamoto Y., Niishi K., Ohkubo I.;

RA RT "Rat tripeptidyl peptidase I: its purification and molecular cloning.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 196-217; 374-392 AND 395-429, AND CHARACTERIZATION.

RC TISSUE-Spleen;

RX MEDLINE=9823562; PubMed=9659384;

RA Vines D.J.; Warburton M.J.;

RT "Purification and characterisation of a tripeptidyl aminopeptidase I

CC activity. May act as a non-specific lysosomal peptidase which generates tripeptides from the breakdown products produced by lysosomal proteinases. Requires substrates with an unsubstituted N-terminus (By similarity).
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide, but also endopeptidase activity.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Activated by autocatalytic proteolytical processing upon acidification (By similarity).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL: AF114167; AAD25043_1; -.

CC DR MEROPS; S53_003; -.

CC KW Glycoprotein; Signal; Lysosome;

CC FT SIGNAL 1 19 BY SIMILARITY.

CC FT PROPEP 20 19 REMOVED IN MATURE FORM (BY SIMILARITY).

CC FT ACT_SITE 196 563 TRIPETIDYL PEPTIDASE I.

CC FT ACT_SITE 360 360 BY SIMILARITY.

CC FT ACT_SITE 475 475 NUCLEOPHILE (BY SIMILARITY).

CC FT ACT_SITE 517 517 BY SIMILARITY.

CC FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).

CC SQ SEQUENCE 563 AA: 61362 MW; 21465RA4C34F4 CRC64;

CC Query Match 90.0% Score 27; DB 1; Length 563;

CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC Qy 1 SVDEY 6

CC Db 274 SLDVEY 279

RESULT 8 TPP1_CANPA STANDARD; PRT; 563 AA.

ID TPP1_CANPA Q9KSB8; AC DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DB Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-I) (Tripeptidyl aminopeptidase) (Lysosomal peptidin insensitive protease) (LPIC).

GN CLN2 OR TPP1.

OS Canis familiaris (Dog).

OC Mammalia; Chordata; Craniata; Vertebrata; Buteleostomi; NCBI_TaxID=9615;

RN 12] RSEQUENCE FROM N.A.

RC Coding sequence and exon/intron organization of the canine CLN2 gene and its exclusion as the locus for cereoid lipofuscinosis in English

RT "Coding sequence and exon/intron organization of the canine CLN2 gene and its exclusion as the locus for cereoid lipofuscinosis in English"

RT Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.

RT "Purification and characterisation of a tripeptidyl aminopeptidase I

RT from rat spleen.";
 RL Biochim. Biophys. Acta 1384:233-242(1998).
 -!- FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I activity. May act as a non-specific lysosomal peptidase which generates tripeptides from the breakdown products produced by lysosomal proteinases. Requires substrates with an unsubstituted N-terminus. Maximum activity at pH 4, unstable above pH 7.
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Activated by autocatalytic proteolytical processing upon acidification (By similarity)
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 553.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AB043870; BAB10570.1;
 KW Hydrolase; Protease; Serine protease; Zymogen; Signal; Lysosome;
 KW Glycoprotein.
 FT SIGNAL 1 19 BY SIMILARITY.
 PROPEP 20 195 REMOVED IN MATURE FORM,
 PT CHAIN 196 563 TRIPETIDYL PEPTIDASE I.
 PT ACT SITE 360 360 BY SIMILARITY .
 PT ACT SITE 475 475 NUCLEOPHILE (BY SIMILARITY) .
 PT ACT SITE 517 517 BY SIMILARITY .
 PT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 PT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 PT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 PT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 PT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 PT CONFLICT 210 210 VG -> SQ (IN REF. 2).
 PT CONFLICT 217 217 GST -> SP (IN REF. 2).
 PT CONFLICT 389 391 61332 MW; B54FPC8625DFEC1 CRC64;
 SEQUENCE 563 AA; 61332 MW; B54FPC8625DFEC1 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 563;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Db 274 SMDVEY 279

Qy 1 SMDVEY 6
 Db 274 SMDVEY 279

RESULT 10
 STF_LAMBD STANDARD; PRT; 774 AA.
 ID P03764; P03745; (Rel. 01, Created)
 AC 016-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Side tail fiber protein.
 GN STP.
 OS Bacteriophage lambda.
 OC Viruses; dsDNA viruses, no RNA stage; Caudoviridae; Siphoviridae;
 OC Lambda-like viruses.

OX NCBI_TAXID=10710;
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=1FO 4570;
 RC MEDLINE=96271012; PubMed=8830045;
 RX RT "Molecular cloning, sequencing, and expression of a cDNA encoding alpha-glucosidase from *Marcus javanicus*.";
 RA Sanger F.; Coulson A.R.; Hong G.F.; Hill D.F.; Petersen G.B.;
 RT "Nucleotide sequence of bacteriophage lambda DNA.";
 BL J. Mol. Biol. 162:729-773 (1982).
 RN IDENTIFICATION AS STP.
 RP MEDLINE=92165720; PubMed=1531648;
 RX Hagaard-Ljungquist E.; Hailing C.; Calendar R.;
 RA

RT "DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages";
 RL J. Bacteriol. 174:1462-1477 (1992).
 CC [3]
 RECONSTRUCTION
 RP MEDLINE=93068310; PubMed=1439823;
 RX

RA Hendrix R.W.; Duda R.L.;
 RT "Bacteriophage lambda PaPa: not the mother of all lambda phages.";
 RL SCIENCE 258:1145-1148 (1992).
 CC -!- MISCELLANEOUS: The common laboratory strain of bacteriophage lambda; lambda PaPa; carries a frameshift mutation relative to Ur-lambda, the original isolate. The Ur-lambda virions have thin, jointed tail fibers (side tail fibers) that are absent from lambda wild type. Relative to lambda PaPa, Ur-lambda has expanded receptor specificity and adsorbs to E.coli cells more rapidly.
 CC -!- SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS CORRECTED IN POSITION TO 396 TO RECREATE THE ORIGINAL STP PROTEIN.
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC DR EMBL; J02459; AAA96555.1; ALT FRAME.
 CC DR EMBL; J02459; AAA96557.1; ALT FRAME.
 CC DR PIR; A04389; Q8BP1L.
 CC DR PIR; A04370; Q8BP2L.
 CC DR InterPro; IPR004089; Chmataxis transd.
 CC DR InterPro; IPR005003; Phage_fiber.
 CC DR Pfam; PF03335; Phage_fiber_6.
 CC KW Fiber protein.
 CC SQ SEQUENCE 774 AA; 77527 MW; CDDDPF85E919123B CRC64;
 CC Query Match 90.0%; Score 27; DB 1; Length 774;
 CC Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC Db 50 SMDVEY 55

RESULT 11
 AGLU_MUCJA STANDARD; PRT; 864 AA.

ID AGLU_MUCJA
 AC 092442;
 AC 092442;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Alpha-glucosidase precursor (BC 3.21.20) (Maltase).
 OS *Mucor javanicus*.
 OC Eukaryota; Fungi; Zygomycetes; Mucorales; Mucoraceae;
 OC Mucor.
 OX NCBI_TAXID=51122;
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=1FO 4570;
 RC MEDLINE=96271012; PubMed=8830045;
 RX RT "Molecular cloning, sequencing, and expression of a cDNA encoding alpha-glucosidase from *Marcus javanicus*.";
 RA Sugimoto M.; Suzuki Y.;
 RT J. Biochem. 119:500-505 (1996)
 RA -!- FUNCTION: HYDROLYSES NOT ONLY MALTO-OLIGOSACCHARIDS BUT ALSO SOLUBLE STARCH.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-linked D-galactose residues with release of D-glucose.
 CC -!- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D67034; BA011053_1; -.
DR InterPro; IPR000322; Glyco_hydro_31.
DR Pfam; PF01055; Glyco_hydro_31_1.
DR PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 1.
DR PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; FALSE_NEG.
KW Hydrolase; Glycoproteine; Glycosidase; Glycoprotein; Signal.
FT SIGNAL_1 22 POTENTIAL.
FT CHAIN 23 864
FT ACT_SITE 430 430 BY SIMILARITY.
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 500 500 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 734 734 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 864 AA; 98761 MW; CPAB4759DC431403 CRC54;

Query Match Score 27; DB 1; Length 864;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDVEY 6
Db 50 SYDVEY 55

RESULT 13
ID MT54_YEAST STANDARD; PRT; 1471 AA.

AC P32492;
ID SPFR_ECOLI STANDARD; PRT; 1120 AA.
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DB Side tail fiber protein homolog from lambdoid prophage Rac.
GN SMFR OR B1372.
OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=562;
RN [1]

SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINEB97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN [2]

SEQUENCE FROM N.A.
STRAIN=K12;
MEDLINEB97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kitakawa M., Mizobuchi K., Mori H., Morii T., Motomura K.,
RA Makino K., Miki T., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Nakade S., Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponds to the 28.0-40.1 min region on the linkage map.";
RL DNA Res 3:363-377(1996).
-1 SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AE00234; AAC74454_1; ALT_INIT.
DR EMBL; D90774; BA014966_1; -.
DR EMBL; D90775; BA014975_1; -.
DR Ecogene; EG13370; STFR.
DR InterPro; IPR004059; Chmtraxis transd.
DR InterPro; IPR005003; Phage_fiber.
DR InterPro; IPR005068; phage_fiber_2.
DR Pfam; PF03335; Phage_fiber_2_1.
DR Pfam; PF03406; Phage_fiber_2_1.
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SEQUENCE 1120 AA; 113779 MW; 542B59D71BE795B4 CRC64;

Query Match Score 27; DB 1; Length 1120;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDVEY 6
Db 50 SYDVEY 55

RESULT 14
ID MT54_YEAST STANDARD; PRT; 1471 AA.

AC P32492;
ID SPFR_ECOLI STANDARD; PRT; 1120 AA.
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DB Myosin-4 isoform.
OS MYO4 OR SHE1 OR YAL029C OR FUN22.
OC Eukaryota; Fungi; Ascomycota; Saccharomyces; Saccharomyces.
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;

RN RP SEQUENCE FROM N.A.
RX MEDLINE:94334370; PubMed=8056830;
RA Haarer B.K., Petzold A., Lillie S.H., Brown S.S.;
RT "Identification of MYO4, a second Class V myosin gene in yeast.";
J. Cell Sci. 107:1055-1064 (1994).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAINS288C / AB972;
RX MEDLINE:952949563; PubMed=7731988;
RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
RA Hall J., Ouellette B.P.P., Keng T., Barton A.B., Su Y., Davies C.K.,
RA Storms R.K.;
RT "The nucleotide sequence of chromosome I from Saccharomyces cerevisiae.";
RT Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813 (1995).

CC -1 FUNCTION: SEEMS TO BE ACCUMULATION IN DAUGHTER CELL OF ASH1, A REPRESSOR OF THE HO ENDONUCLEASE.
CC -1 SUBCELLULAR LOCATION: ACCUMULATES PREFERENTIALLY IN GROWING BUDS.
CC -1 SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1 SIMILARITY: CONTAINS 1 DILUT DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M90057; AAC37409.1; -.
 DR EMBL; U12980; AAC05003.1; -.
 DR PIR; S30794; S30790.
 DR HSSP; P10581; 1SR2.
 DR SGD; S0000027; MYO4.
 DR InterPro; IPR00210; DIL.
 DR InterPro; IPR00048; IQ_region.
 DR InterPro; IPR00009; Myosin_N.
 DR InterPro; IPR001609; myosin_head.
 PFam; PF00063; myosin_head_1.
 PFam; PF00512; IQ_5.
 DR Pfam; PF01843; DIV_1.
 DR Pfam; PF02736; Myosin_N_1.
 PRINTS; PR00193; MYOSINHEAVY.
 DR Prodom; PD000355; myosin_head_1.
 DR Prodom; PD003376; DIV_1.
 SMART; SM00015; IQ_2.
 SMART; SM00242; MYSC_1.
 PROSITE; PS50096; IQI_2.
 Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 KW Coiled coil; Alkylation.
 FT DOMAIN 1 777 MYOSIN HEAD-LIKE.
 FT DOMAIN 876 898 IQ_1.
 FT DOMAIN 899 928 IQ_2.
 FT DOMAIN 938 1063 COILED COIL.
 FT DOMAIN 1064 1471 CARBOXYL TERMINAL.
 NP BIND 165 172 ATP (POTENTIAL).

Qy	1	SYDNEY	6
Query Match	90.0%	Score 27:	DB 1;
Best Local Similarity	83.3%	Pred. No.	2.3e+02;
Matches 5;	Conservative	1;	Mismatches 0;
			Indels 0;

RESULT 14			
POLG-PRSVH	STANDARD;	PRT;	3344 AA.
ID _PRSVH			
AC Q01901;			
DT 01-OCT-1993	(Rel. 27, Created)		
DT 01-OCT-1996	(Rel. 34, Last sequence update)		
DT 15-JUN-2002	(Rel. 41, Last annotation update)		
DE Genome polyprotein	(Contains: N-terminal protein (P1); Helper		
DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa			
DE protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2 (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A); NI(A) (EC 3.4.22.44) (49 kDa proteinase); 49 kDa Pro); Nuclear			
DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP).			
DS Papaya ringspot virus (strain P / mutant HA).			
OS Virus; ssRNA positive-strand viruses, no DNA stage; Potyviridae;			
OC			

OX NCBI_TAXID:31731;
 RN [1]
 SEQUENCE FROM N.A.
 RP Wang C.H., Bau H.J., Yeh S.D.;
 RA "Comparison of the nuclear inclusion b protein and coat protein genes
 RT of five papaya ringspot virus strains distinct in geographic origin
 RT and pathogenicity.";
 RT Phytopathology 84:1205-1210(1994).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93019006; PubMed=1402799;
 RX MEDLINE=93019006;
 RA Yeh S.D., Jan P.J., Chiang C.H., Doong T.J., Chen M.C.,
 RA Chung P.H., Bau H.J.;
 RT "Complete nucleotide sequence and genetic organization of papaya
 RT ringspot virus RNA";
 RT J Gen Virol 73:2531-2541(1992)
 PI

[3] IRN
SEQUENCE OF 2561-3344 FROM N.A.
MEDLINE=9309098; PubMed=1456898;
Wang C.H., Yeh S.D.;
"Nucleotide sequence comparison of the 3'-terminal regions of severe, mild, and non-pathogenic strains of papaya ringspot virus.",
Arch. Virol. 127:345-354 (1992).

-1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.

-1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.

-1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.

-1- CATALYTIC ACTIVITY: Hydrolyzes glutaminyl bonds, and activity is further restricted by preferences for the amino acids in P6 - P1, that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-Gln (Ser or Gly) for the enzyme from tobacco etch virus. The natural substrate is the viral polyprotein, but other proteins and oligopeptides containing the appropriate consensus sequence are also cleaved.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N' diphosphate + [RNA] (N).

-1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-[Gly bond at its own C-terminus commonly in the sequence -Tyr-Xaa-Val-Gly-]-Gly, in the processing of the potyviral polyprotein.

-1- PTM: VPg IS COVALENTLY LINKED TO THE GENOMIC RNA.

-1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.

-1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.

-1- SIMILARITY: NI1 A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.

-1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

EMBP; X67673; CAA47905; 1; -.
EMBL; S46722; AA823789; 1; -.
EMBL; X67672; CAA47904; 1; -.
PIR; S24785; S24785.
MEROPS; C04_009; -.
MEROPS; C06_001; -.
MEROPS; S30_001; -.
InterPro; IPR001410; DBAD.
InterPro; IPR01650; Helicase_C.
InterPro; IPR001730; Peptidase_C4.
InterPro; IPR001456; Peptidase_C6.
InterPro; IPR002540; Poty_P1.
InterPro; IPR001592; Poty_coat.
InterPro; IPR001205; RNA_Pol_P3D.
InterPro; IPR001554; Ser_protease_TRY.
PFAM; PF000851; helicase_C; 1.
PFAM; PF000680; RNA_dep_RNA_pol; 1.
PFAM; PF00767; Poty_coat; 1.
PFAM; PF000851; Peptidase_C6; 1.
PFAM; PF000853; Peptidase_e-C4; 1.
PFAM; PF01577; Poty_P1; 1.
PRINTS; PR00066; NIAPOTYPPASE.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELICE; 1.
Hydrolase; Transferase; Thiol Protease; RNA-directed RNA Polymerase; Coat protein; Polyprotein; Covalent protein-RNA linkage; ATP-binding.
CHAIN 1 529 N-TERMINAL PROTEIN (BY SIMILARITY).
CHAIN 530 1149 HELPER COMPONENT PROTEINASE (BY SIMILARITY).
FPT 1150 ? PROTEIN P3 (BY SIMILARITY).
FPT 1151 ? 6 KDA PROTEIN 1 (BY SIMILARITY).
FPT 1152 ? PROTEIN P3 (BY SIMILARITY).

FT CHAIN 1402 2036 CYTOPLASMIC INCLUSION PROTEIN (BY SIMILARITY).
 FT CHAIN 2037 2093 6 KDA PROTEIN 2 (BY SIMILARITY).
 FT CHAIN ? 2520 GENOM-LINKED PROTEIN (BY SIMILARITY).
 FT CHAIN ? NUCLEAR INCLUSION PROTEIN A (BY SIMILARITY).
 FT CHAIN 2521 3037 NUCLEAR INCLUSION PROTEIN B (BY SIMILARITY).
 FT CHAIN 3038 3344 COVALENT LINKAGE OF VIRAL RNA (BY SIMILARITY).
 FT BINDING 2156 2156 ATP (POTENTIAL).
 FT NP BIND 1486 1493 SEQUENCE 3344 AA; 381040 MW; E90CD7523AC5243D CRC64;
 Query Match 90.0%; Score 27; DB 1; Length 3344;
 Best Local Similarity 83.33%; Pred. No. 5.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SVDEY 6
 Db 149 SVDEY 154

RESULT 15
 YOFP_ECOLI STANDARD PRT; 47 AA.
 AC P72265;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DS Hypothetical protein yobF.
 OS YOFP OR B1824 OR Z2869 OR ECSC2534.
 OS Escherichia coli and Escherichia coli Q157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NCBI_TaxID=562, 83334;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426517; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Burland V., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=0157:H7 / ATCC 700927;
 RX MEDLINE=2107435; PubMed=1120551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimmlanta B.T., Poramasis K.,
 RA Apodaca J., Amanataraman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=0157:H7 / RIMD 050952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo B., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Title: US-09-919-703-1
Exact score: 30
Sequence: 1 SVDEY 6
Scoring table: BLOSUM62
Copyright GenCore version 5.1.3
(c) 1993 - 2003 Compugen Ltd.
4 protein - protein search, using sw model
Run on: January 10, 2003, 11:45:36 : Search time 15 Seconds
(without alignments)
38.454 Million cell updates/sec
probable tail fiber
probable tail fiber
probable tail fiber
probable tail fiber
hypothetical prote
probable tail comp
probable ABC trans
ABC transporter A
hypothetical prote
alpha-glucosidase
probable tail fiber
probable membrane
peptidase related
probable tail fiber
myosin MYO4 - y
genome polyprotein
D90734
C90769
E85816
A85719
A85741
F95591
HB1971
C81026
E96830
JC4624
B90835
C55693
D90387
G64887
S30790
J01899
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

ALIGNMENT S.

SUMMARIES									
Result No.	Score	Query Length	DB ID	Description					
1	30	100.0	128	2	streptokinase A (B				
2	30	100.0	157	2	arenate reductase				
3	30	100.0	279	2	probable lectin 2				
4	30	100.0	415	1	BZSO				
5	30	100.0	440	1	streptokinase (EC				
6	30	100.0	440	2	streptokinase prec				
7	30	100.0	440	2	streptokinase G pr				
8	30	100.0	1274	2	transferrin-like p				
9	29	96.7	186	2	hypothetical Prote				
10	29	96.7	215	2	probable phosphoProte				
11	29	96.7	1207	2	hypothetical prote				
12	29	96.7	1595	2	endo-1,4-beta-xyla				
13	28	93.3	311	1	primosome component				
14	28	93.3	629	2	hypothetical Prote				
15	27	90.0	50	2	ma for outer membra				
16	27	90.0	180	2	protein B0244.9 [1]				
17	27	90.0	271	2	probable tail fiber				
18	27	90.0	271	2	hypothetical Prote				
19	27	90.0	368	2	outer membrane 40K				
20	27	90.0	368	2	porin precursor -				
21	27	90.0	371	2	porin fona precurs				
22	27	90.0	375	2	hypothetical Prote				
23	27	90.0	391	2	probable transcript				
24	27	90.0	399	2	probable PPG Prote				
25	27	90.0	401	1	hypothetical Prote				
26	27	90.0	407	2	probable tail fiber				
27	27	90.0	437	2	probable tail fiber				
28	27	90.0	437	2	probable tail fiber				
29	27	90.0	437	2	probable tail fiber				

A;Residues: 1-157 <KUR>
A;Cross-references: GB:AB008917; PIDN:AAL52851.1; PID:917983693; GSDB:GN00190
A;Experimental source: strain 16M
A;Gene: BME11670
A;Map position: I
Query Match 100.0%; Score 30; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 114 SVDEY 119

RESULT 3

T09620

probable lectin 2 precursor - alfalfa

C;Species: *Medicago sativa* (alfalfa)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

C;Accession: T09620

R;Brill, L.M.; Pieterrel, V.R.

Submitted to the EMBL Data Library, March 1998

A;Description: Legume seed lectin genes: sequence of Mslec2 from Alfalfa, Alfalfa and Sk

A;Reference number: Z16780

A;Status: Preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: 1-279

A;Cross-references: EMBL:Y16754

A;Experimental source: cultivar Chief

C;Genetics:

A;Gene: lec2

C;Superfamily: plant lectin

C;Keywords: calcium, glycoprotein, lectin

F;1-25/Domain: signal sequence #status predicted <SIG>

F;26-279/Product: probable lectin 2 #status predicted <MAT>

Query Match 100.0%; Score 30; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 157 SVDEY 162

RESULT 6

S02723

streptokinase G precursor - *Streptococcus* sp. (strain 19908)C;Species: *Streptococcus* sp.

C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 22-Jun-1999

C;Accession: S02723

R;Walter, F.; Siegel, M.; Malke, H.

Nucleic Acids Res 17, 1262, 1989

A;Title: Nucleotide sequence of the streptokinase gene from a group-G *Streptococcus*.

A;Reference number: S02723; PMID:89160265; PMID:2922269

A;Cross-references: EMBL:X13400; NID:947095; PIDN:CAA31766.1; PID:947096

A;Molecule type: DNA

A;Residues: 1-440 <WAL>

C;Genetics:

A;Gene: skg

C;Superfamily: streptokinase

C;Domain: signal sequence #status predicted <MAT>

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-440/Product: streptokinase #status predicted <MAT>

Query Match 100.0%; Score 30; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 183 SVDEY 188

RESULT 7

S02724

streptokinase A precursor - *Streptococcus* pyogenes (strain SF130/13)C;Species: *Streptococcus* pyogenes

C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 22-Jun-1999

C;Accession: S02724

R;Walter, F.; Siegel, M.; Malke, H.

Nucleic Acids Res 17, 1261, 1989

A;Title: Nucleotide sequence of the streptokinase gene from a *Streptococcus pyogenes* t-

A;Reference number: S02724; PMID:89160264; PMID:2646590

A;Cross-references: EMBL:S02224

A;Residues: 1-440 <WAL>

A;Cross-references: EMBL:X13399; NID:947435; PIDN:CAA31765.1; PID:947436

Query Match 100.0%; Score 30; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 157 SVDEY 162

Query Match 100.0%; Score 30; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 208 SVDEY 213

Query Match 100.0%; Score 30; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 183 SVDEY 188

RESULT 4

B280

streptokinase (EC 3.4.-.-) - *Streptococcus* sp.C;Species: *Streptococcus* sp.

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 23-Aug-1996

C;Accession: A00967

R;Jackson, K.W.; Tang, J.

Biochemistry 21, 6620-6625, 1982

A;Title: Complete amino acid sequence of streptokinase and its homology with serine protease

A;Reference number: A00967; PMID:7760891

A;Cross-references: EMBL:A00967

A;Molecule type: protein

A;Residues: 1-415 <JAC>

A;Note: 169-Asp and 181-Asp were also found

A;Note: this protein is not a protease, but it activates plasminogen by complexing with

C;Superfamily: streptokinase

C;Keywords: hydrolase

A;Cross-references: EMBL:X13399; NID:947435; PIDN:CAA31765.1; PID:947436

Query Match 100.0%; Score 30; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDEY 6
Db 157 SVDEY 162

C;Genetics:
 A;Gene: ska
 C;Superfamily: streptokinase
 F;1-26/Domain: signal sequence #status predicted <SIC>
 F;27-440/Product: streptokinase #status predicted <NBT>
 Query Match 100.0% Score 30; DB 2; Length 440;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SYDVEY 6
 Db 183 SYDVEY 188

RESULT 8
 T10729

transferrin-like protein Tf-1, salt-induced - green alga (*Dunaliella salina*)
 C;Species: Dunaliella salina

C;Accession: T10729
 C;Description: submitted to the EMBL Data Library, November 1996
 A;Reference number: Z17101
 A;Accession: T10729

A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1274 <FIG>
 A;Cross-references: EMBL:U77059; NID:91684791; PID:91684792

C;Genetics:
 A;Gene: ttfl

C;Superfamily: transferrin repeat homology
 Query Match 100.0% Score 30; DB 2; Length 1274;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SYDVEY 6
 Db 1186 SYDVEY 1191

RESULT 9
 T31951

hypothetical protein B0047.4 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Accession: T31951
 C;Description: The sequence of C. elegans cosmid B0047.

A;Reference number: Z21102
 A;Accession: T31951
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-186 <PAU>

A;Cross-references: EMBL:AF016658; PIDN:AB066044.1; GSPDB:GN00020; CESP:B0047.4

C;Genetics:
 A;Gene: CESP_B0047.4

A;Map position: 2
 A;Introns: 51/2

Query Match 96.7% Score 29; DB 2; Length 186;
 Best Local Similarity 83.3%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SYDVEY 6
 Db 79 SYDVEY 84

RESULT 10
 T31082

endo-1,4-beta-xylanase (EC 3.2.1.8) - Caldicellulosiruptor sp.
 C;Species: Caldicellulosiruptor sp.
 C;Accession: T31082
 C;Description: Family 10 and 11 xylanase genes from Caldicellulosiruptor sp. Rtc69B.1.

A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1595 <MOR>

A;Cross-references: EMBL:AF036923; NID:92760904; PID:92760908; PIDN:AA095325.1

C97297 probable phosphatase, HAD superfamily [imported] - Clostridium acetobutylicum
 C;Species: Clostridium acetobutylicum
 C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001
 C;Accession: C97297
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cj
 A;Reference number: A96900; PMID:21359325; PMID:21359325
 A;Accession: C97297
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-215 <KUR>
 A;Cross-references: GB:AB001437; PIDN:AAK81166.1; PID:915026304; GSPDB:GN00168
 A;Experimental source: Clostridium acetobutylicum ATCC824
 A;Gene: CAC12331
 C;Genetics:

Query Match 96.7% Score 29; DB 2; Length 215;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SYDVEY 6
 Db 25 SIDVEY 30

RESULT 11
 T16011

hypothetical protein F09F7.3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Accession: T16011
 C;Description: The sequence of C. elegans revision 20-Sep-1999
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1207 <PAU>

A;Cross-references: EMBL:U00050; NID:9485108; PID:9485110; PIDN:AA050695.1; CBSP:F09F7.
 C;Genetics:

A;Gene: CBSP:F09F7.3
 A;Introns: 23/1-297/3; 600/2; 630/2; 724/3; 789/3; 916/1; 1102/2; 1150/3
 C;Superfamily: DNA-directed RNA Polymerase 132K polypeptide
 Query Match 96.7% Score 29; DB 2; Length 1207;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SYDVEY 6
 Db 170 SIDVEY 175

RESULT 12
 T31082

endo-1,4-beta-xylanase (EC 3.2.1.8) - Caldicellulosiruptor sp.
 C;Species: Caldicellulosiruptor sp.
 C;Accession: T31082
 C;Description: Family 10 and 11 xylanase genes from Caldicellulosiruptor sp. Rtc69B.1.

A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1595 <MOR>

A;Cross-references: EMBL:AF036923; NID:92760904; PID:92760908; PIDN:AA095325.1

C:Genetics:
A:Gene: xynB
C:Keywords: glycosidase; hydrolase

Query Match 96.7%; Score 29; DB 2; Length 1595;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 1548 SIDVEY 1553

RESULT 13

IQB54 Primosome component (helicase loader) dnaI - *Bacillus subtilis*
N;Alternate names: dnaA protein homolog, 44K; hypothetical protein Y (dnab 3' region)
C;Species: *Bacillus subtilis*
C;Accession: B24720; C26580; F69617
R;Ogasawara, N.; Moriya, S.; Maza, P.G.; Yoshikawa, H.
Nucleic Acids Res. 14, 9889-9899, 1986
A;Title: Nucleotide sequence and organization of dnab gene and neighbouring genes on the
A;Reference number: A93650; MUID:87117549; PMID:3027671
A;Accession: B24720
A;Molecule type: DNA
A;Residues: 1-11 <OGA>
A;Cross-references: GB:X04963; NID:939880; PIDN:CAA28633.1; PID:939881
R;Hoshino, T.; McKenzie, T.; Schmidt, S.; Tanaka, T.; Sueoka, N.
Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987
A;Title: Nucleotide sequence of *Bacillus subtilis* dnaB: a gene essential for DNA replica
A;Reference number: A94709; MUID:87118226; PMID:3027897
A;Accession: C26580
A;Molecule type: DNA
A;Residues: 1-18, 'N' 20-23, 'T' 25-206 <HOS>
R;Kunst, F.; Ogasawara, N.; Moszter, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 330, 249-256, 1997
A;Authors: Poulier, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallereich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hosappel, S.; Hull, M.P.; Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel, Y. M.; Ogawa, K.; Ogiwara, A.; Onega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Satoh, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Seikowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshihawa, H.F.; Zumsteg, B.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Accession number: A69580; MUID:98044033; PMID:9384377

A;Status: nucleic acid sequence not shown; translation not shown

A;Residues: 1-311 <RUN>

A;Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14858.1; PID:g2635363

A;Experimental source: strain 168

C;Genetics:

A;Gene: dnaI
C;Superfamily: 44K dnaA protein homolog
C;Keywords: ATP; nucleotide binding; P-loop
P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

T27619 hypothetical protein ZC504.2 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Accession: T27619

R;Kershaw, J.

Submitted to the EMBL Data Library, July 1995

A;Reference number: Z20394

A;Status: Preliminary; translated from GB/EMBL/DDJB

A;Accession: T27619

A;Molecule type: DNA

A;Residues: 1-629 <WIL>

A;Cross-references: EMBL:Z50029; PIDN:CAA90341.1; GSFPDB:GN00028; CBSP:ZC504.2

A;Experimental source: clone ZC504

A;Genetics:

A;Gene: CESP:ZC504.2

A;Map position: X

A;Introns: 11/3; 63/1; 114/3; 166/2; 431/2; 405/1; 431/2; 405/2;

C;Superfamily: acetylcholine receptor

Query Match 93.3%; Score 28; DB 2; Length 629;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Species: Fusobacterium nucleatum (strain FeVI) (fragment)

C;Accession: 30-Sep-1993 #sequence_revision 30-Sep-1993

C;Title: Fusobacterium nucleatum

A60745 major outer membrane protein - *Fusobacterium nucleatum*

C;Accession: 30-Sep-1993 #sequence_revision 30-Sep-1993

C;Title: Fusobacterium nucleatum

A60745 major outer membrane protein - *Fusobacterium nucleatum* (strain FeVI) (fragment)

C;Accession: A60745

R;BAKKEN, V.; AARO, S.; JENSEN, H.B.

J. Gen. Microbiol. 135, 3253-3262, 1989

A;Title: Purification and partial characterization of a major outer-membrane protein of

A;Authors: Fouiliger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallo-

reich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hosappel, S.; Hosono, S.; Hull, M.P.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel,

Y. M.; Ogawa, K.; Ogiwara, A.; Onega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,

Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Satoh, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Seikowska, A.; Seron,

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Accession number: A69580; MUID:98044033; PMID:9384377

A;Status: nucleic acid sequence not shown; translation not shown

A;Residues: 1-311 <RUN>

A;Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14858.1; PID:g2635363

A;Experimental source: strain 168

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

P;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match 93.3%; Score 28; DB 1; Length 311

Best Local Similarity 66.7%; Pred. No. 57;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

C;Genetics

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using bw model

Run on: January 10, 2003, 11:45:21 ; Search time 29 Seconds

(without alignments)

42,630 Million cell updates/sec

Title: US-09-919-703-1

Perfect score: 30

Sequence: 1 SVDBY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:

671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	3.0	100.0	157	16	Q8YF57	SEQUENCE FROM N.A. STRAIN 16M / ATCC 23456 / BIOTYPE 1; MEDLINE=20020109; PubMed=11756688;
2	3.0	100.0	279	10	Q49899	DeVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T., Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G., Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Gotsman E., Selkov E., Elzer P.H., Hugus S., O'Callaghan D., Letesson J.-J., RA Haselkorn R., Kyriakis N., Overbeek R.
3	3.0	100.0	414	2	Q53284	"The genome sequence of the facultative intracellular pathogen Brucella melitensis."
4	3.0	100.0	891	5	Q25177	RT Brucella melitensis." RT
5	3.0	100.0	1274	10	P93125	RJL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
6	2.9	96.7	60	5	Q9NFR4	DR EMBL; AB009602; AAL52851.; DR InterPro; IPR00106; Low_mwt_PTPase.
7	2.9	96.7	74	16	Q9LSP9	DR PF01451; LMWPC; 1. DR SMART; SM00226; LMWPC; 1. DR Complete proteome; KW SQ Sequence 157 AA; 17446 MW; D8AA7748107ECB65 CRC64;
8	2.9	96.7	186	5	Q16567	DR Best Local Similarity 100.0%; Pred. No. 26; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
9	2.9	96.7	215	16	Q9T784	DR Query Match 100.0%; Score 30; DB 16; Length 157;
10	2.9	96.7	520	5	Q8W5V6	DR Q9wsf9 drosophila
11	2.9	96.7	547	5	Q8WSF9	DR Q9wsf9 aplasia californica
12	2.9	96.7	615	3	Q9P861	DR Q9P861 arcula adenoides
13	2.9	96.7	742	5	Q8ZV19	DR Q9svv9 drosophila
14	2.9	96.7	830	2	Q9AGFB	DR Q9gq8 arthropods
15	2.9	96.7	836	10	Q9PGNO	DR Q9gn0 arabidopsis
16	2.9	96.7	1154	5	Q27492	DR Q27492 caenorhabditis

RESULT 2		Query Match		Score 30; DB 2;		Length 414;	
Q49899		Best Local Similarity 100.0%;		Pred. No. 71;			
ID	PRT;	Matches 6;		Mismatches 0;		Indels 0;	Gaps 0;
AC	PRELIMINARY;						
DT	04-JUN-1998 (TREMBLrel. 06, Created)						
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)						
DR	01-MAR-2002 (TREMBLrel. 20, Last annotation update)						
DE	Lectin precursor.						
GN	LBG2.						
OS	Medicago sativa (Alfalfa).						
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;						
OC	euksorids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Madicago.						
OX	NCBI_TaxID=3879;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	SPRAIN=CV. CHIEP;						
RA	Brill L.M., Pieterse V.R.;						
RT	"Legume seed lectin genes: sequence of Ms1ec2 from Alfalfa, Alfalfa and Sweet Clover southern blot analysis, and stable transformation of Alfalfa with antisense-lectin constructs."						
RL	Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.						
DR	EMB: Y16754; CAA76366 1; -.						
DR	HSSP; P04122; ILOB.						
DR	InterPro; IPR00085; Lectin_legA.						
DR	InterPro; IPR001220; Lectin_legB.						
PFAM	PF00138; lectin_legA_1.						
PFAM	PF00139; lectin_legB_1.						
DR	ProDom; P000671; Lectin_legA_1.						
DR	ProDom; P000711; Lectin_legB_1.						
DR	PROSITE; PS00308; LECTIN_LIGUME_ALPHA_1.						
DR	PROSITE; PS00307; LECTIN_LIGUME_BETA; UNKNOWN_1.						
KW	Signal.						
FT	SIGNAL 1 25 POTENTIAL.						
FT	CHAIN 26 279 LECTIN.						
SEQUENCE	279 AA; 30551 MW; 1315F022BABDA360 CRC64;						
Query Match	Score 30; DB 10; Length 279;						
Best Local Similarity 100.0%;	Pred. No. 47;						
Matches 6; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
Qy	1 SVDTVEY 6						
Db	208 SVDTVEY 213						
RESULT 3		Query Match		Score 30; DB 5;		Length 891;	
QS3284	ID	PRELIMINARY;	PRT;	414 AA.			
AC	QS3284						
DT	01-NOV-1996 (TREMBLrel. 01, Created)						
DR	01-NOV-1996 (TREMBLrel. 01, Last sequence update)						
DR	01-MAR-2002 (TREMBLrel. 20, Last annotation update)						
DE	SKC-2.						
OS	Streptococcus equisimilis.						
OC	Bacteria; Firmicutes; Bacilli; Clostridium group; Lactobacillales;						
OC	Streptococcaceae; Streptococcus.						
NCBI_TaxID=119602;							
RN	[1]						
RP	SEQUENCE FROM N.A.						
MEIDLINE=93000998; PubMed=1368792;							
RA	Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,						
RA	Rubiera R., Pedraza A., Padron G., Antuch W., de la Fuente J.,						
RA	Herrera L.; "High level expression of streprokinase in Escherichia coli."						
RL	Biotechnology 10:1138-1142 (1992).						
DR	InterPro; IPR004093; Streplokinase.						
DR	PFAM; PF02822; Staphylokinase; 3.						
SEQUENCE	414 AA; 47254 MW; F75B85831B766904 CRC64;						
RESULT 4		Query Match		Score 30; DB 1;		Length 891;	
Q52177	ID	PRELIMINARY;	PRT;	891 AA.			
AC							
DT	01-NOV-1996 (TREMBLrel. 01, Created)						
DR	01-NOV-1996 (TREMBLrel. 01, Last sequence update)						
DR	01-JUN-2002 (TREMBLrel. 21, Last annotation update)						
DE	Receptor protein-tyrosine kinase (Fragment).						
OS	Hydra attenuata (Hydra) (Hydra vulgaris).						
OC	Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;						
OC	Hydridae; Hydra.						
NCBI_TaxID=6087;							
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=NCTC 111.						
RA	Chen R.L., Steele R.E.						
RL	Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.						
DR	AA65223.1; -.						
DR	HSSP; P1132; 1FGK.						
DR	InterPro; IPR002106; AACNA_ligaseII.						
DR	InterPro; IPR000719; Euk_pk kinase.						
DR	InterPro; IPR003160; Ig_Like.						
DR	InterPro; IPR001006; Ig_XHC.						
DR	InterPro; IPR001245; Tyr_pk kinase.						
DR	PFam; PF00047; Ig_2.						
DR	PFam; PF00069; pk kinase; 1.						
DR	PRINTS; PRO0109; TYRKINASE.						
DR	ProDom; PD000001; Euk_pk kinase; 1.						
DR	SMART; SM00410; Ig_Like; 2.						
DR	PROSITE; PS003339; AA_TRNA_LIGASE_I_2; UNKNOWN_1.						
DR	PROSITE; PS50011; PROTEIN_KINASE_DON; 1.						
DR	PROSITE; PS00103; PROTEIN_KINASE_TYR; 1.						
KW	Kinase: Tyrosine-protein Kinase.						
FT	NON_TER.						
SEQ	891 AA; 100976 MW; 1CFDF4BCCA298176 CRC64;						
RESULT 5		Query Match		Score 30; DB 5;		Length 891;	
Q53125	ID	PRELIMINARY;	PRT;	1274 AA.			
AC							
DT	01-MAY-1997 (TREMBLrel. 03, Created)						
DR	01-MAY-1997 (TREMBLrel. 03, Last sequence update)						
DR	01-JUN-2002 (TREMBLrel. 21, Last annotation update)						
DE	Tff-1.						
GN	Dunaliella salina.						
OC	Bukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;						
OC	Dunaliellaceae; Dunaliella.						
NCBI_TaxID=3046;	[1]						
RP	SEQUENCE FROM N.A.						
RA	Fisher M., Gokhman I., Pick U., Zamir A.;						
RA	Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.						
DR	EMBL; U77059; AAB36531.1; -.						

Waterson R.; Submitted (JUL-1997) to the EMBL/GenBank/DDJB databases.
 InterPro; IPR02083; MATH.
 Pfam; PF00917; MATH; 1;
 SMART; SM00061; MATH; 1;
 SEQUENCE 186 AA; 21527 MW; 9E216690AB0E1469 CRC64;

Query Match 96.7%; Score 29; DB 5; Length 186;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

' 1 SIDVY 6
 : | :| |
 . 79 SIDVY 84

SULT 9

, Q97B84 PRELIMINARY; PRT; 215 AA.
 , 01-OCT-2001 (TREMBLrel. 18, Created)
 , 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 , 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 ; Predicted phosphatase, HAD superfamily.
 CAC231.

SEQUENCE FROM N.A.
 Bacteria; Firmicutes; Clostridium group; Clostridia;
 Clostridiales; Clostridiaceae; Clostridium.
 [1] _TaxID=1488;

STRAN=ATC 824 / DSM 792 / VKM B-1787;
 MEDLINE=21359325; PubMed=11456286;
 Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 Gibson R., Lee H.M., Dubois J., Olu D., Hitti J., Wolf Y.I.,
 Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 Bennett G.N., Koornin B.V., Smith D.R.;
 "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum.";
 J. Bacteriol. 183:4823-4838 (2001).
 InterPro; IPR001454; Hydrolase; 1;
 Pfam; PF00702; Hydrolase; 1;
 Complete proteome.
 SEQUENCE 215 AA; 24421 MW; C4334F6313CBPC15 CRC64;

Query Match 96.7%; Score 29; DB 16; Length 215;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

' 1 SIDVY 6
 : | :| |
 . 25 SIDVY 30

SULT 10

, Q9W5V6 PRELIMINARY; PRT; 520 AA.
 , 09W5V6; Q9W5V5;
 , 01-MAY-2000 (TREMBLrel. 13, Created)
 , 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 ; Putative nicotinic acetylcholine receptor alpha 80B (Fragment).
 NACR-ALPHA 80B OR CG1752.
 Drosophila melanogaster (Fruit fly).
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Drosophila;
 Ephdroidea; Prosophaillidae; Drosophilidae.
 [1] _TaxID=7227;
 SEQUENCE FROM N.A.

STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Adams M.D., Celniker S.E., Gocayne J.D., Holt R.A., Galle R.P.,
 RA Amanatides P.G., Scheerer S.B., Li P.W., Roskies R.A., Galle R.P.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton D.C., Rogers Y.-H.C., Blazej R.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews P., Pfannkoch C., Baldwin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.M.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew J., Dietz S.M.,
 RA Dodge K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoertin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A.,
 RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kettman K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPhereson D.,
 RA Merkulov G., Milashina N.V., Molaberry C., Morris J., Mosheri A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muniz D.M., Nelson D.L.,
 RA Nelson D.R., Nixon K., Nixon K., Nursultan D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou N., Zhou X., Zhu S.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Rubin J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).

-1 - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1 - SIMILARITY: AE002665; AAP45409.1; -.

CC DR -1 - SIMILARITY: AE002665; AAP45409.1; -.

CC DR InterPro; IPR001175; GABA_A receptor.
 DR InterPro; IPR001175; Neur_Channel.
 DR InterPro; IPR001175; Neur_Channel.
 DR PFM; PF02931; Neur_chan_BDD; 1.
 DR PFM; PF02922; Neur_chan_memb; 1.
 DR PRINTS; PRO00252; NRIONCHANNEL.
 DR PROSITE; PS00236; NEUROTRION CHANNEL; 2.
 DR Glycoprotein; Hypothetical protein; Ionic channel;
 KW Post-synaptic membrane; Transmembrane.
 FT N-TER 1 1
 SQ SEQUENCE 520 AA; 59339 MW; 5B2AE734E78E7D4D CRC64;

Query Match 96.7%; Score 29; DB 5; Length 520;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RBSLT 11

QY 1 SIDVY 6
 ID Q9W5F9 PRELIMINARY;
 AC Q9W5F9; PRELIMINARY;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DS Non-alpha nicotinic acetylcholine receptor subunit
 OS Aplysia californica (California sea hare).

OC Bukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidea;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;

[1]

RN SEQUENCE FROM N.A.;
RP "Meleshkevich E.A., Morez L.L." -
SAZAEV R.I., Pachin Y.V., Meleshkevich E.A., Morez L.L.;
RT "Non-alpha nicotinic acetylcholine receptor subunit from the
individual identified neurons of Aplysia californica.";
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AP02876; AP037250-1;
DR InterPro; IPR000198; GABA receptor.
DR InterPro; IPR001175; Neur_channel.
PFam; PF02931; Neur_chan_LBP; 1.
PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00880; LINC; 1.
DR PROSITE; PS00236; NEUROTRION_CHANNEL; UNKNOWN_1.
KW RESEQUENCE.

SEQUENCE 547 AA; 62535 MW; 3D3F011D78605C91 CRC64;

Query Match Score 96.7%; Score 29; DB 5; Length 547;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

Db 155 SIVEY 160

RESULT 12

O98861 ID Q98861 PRELIMINARY;

AC Q98861;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DB Ferro-O2-oxidoreductase precursor.

GN AFET3.

OS Arxula adeninivorans.

OC BUKARYOTA; Fungi; Ascomycota; Saccharomycetes; Saccharomyces; Saccharomyces; Dipodascaceae; Arxula.

NCBI_TaxID=37620;

[1]

RP SEQUENCE FROM N.A.

RA Wartmann T., Stephan U.W., Bube I., Boer E., Melzer M.,
Manteuffel R., Stoltenburg R., Kunze G.;
RT "Post-translational modifications of the ABP73 gene product - a
component of iron transport system in budding cells and mycelia of the
yeast Arxula adeninivorans LS3.";
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AAC7783; CB90817.1; -.

HSSP; PS7064; 1AOZ.

InterPro; IPR01111; Cu-oxidase.

DR InterPro; IPR00235; Multicopper oxidase2.

PFam; PF0394; Cu-oxidase; 3.

PROSITE; PS00079; MULTICOPPER_OXIDASE2; 1.

KW SIGNAL.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 615 AA; 69624 MW; 79200D42A28715E4 CRC64;

SQ SEQUENCEB. 615 AA; 69624 MW; 79200D42A28715E4 CRC64;

Query Match Score 96.7%; Score 29; DB 3; Length 615;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

Db 311 SVDIEY 316

RESULT 13

OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyrdoidea; Drosophilidae; Drosophila.

[1]

RP SEQUENCE FROM N.A.

DR STRAIN=BERKELEY;

DR Stapleton M., Brookstein P., Hong L., Agbayani A., Carlson J.,

DR Champ M., Chaves C., Dorsett V., Parfan D., Frise E., George R.,

DR Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

DR Nunoo J., Pacleb J., Paregas V., Park S., Phouanavong S., Wan K.,

DR Yu C., Lewis S.E., Rubin G.M., Celiker S.;

DR Submitted (DBC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY069758; AAL39903; 1; -.

DR SEQUENCE 742 AA; 84075 MW; 85998BF365194B39 CRC64;

RESULT 15

O9FCN0	PRELIMINARY;	PRT;	836 AA.
ID O9FCN0			
AC Q9FCN0;			
DT 01-MAR-2001	(TREMBrel. 16, Created)		
DT 01-MAR-2001	(TREMBrel. 16, Last sequence update)		
DT 01-DBC-2001	(TREMBrel. 19, Last annotation update)		
DE Similarity to unknown protein.			
OS Arabidopsis thaliana (Mouse-ear cress).			
OC Eukaryota; Viridiplantae; Streptophytina; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
NCBI_TaxID=3702;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=COLUMBIA;			
RA Kaneko T., Katoch T., Asamizu S., Sato S., Nakamura Y., Kotani H.,			
RA Tabata S.;			
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";			
RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.			
DR EMBL: A0025621; BR09754.1;			
SQ SEQUENCE 836 AA; 92259 MW; A875BAE2P18F61BD CRC64;			
Query Match 96.7%; Score 29; DB 10; Length 836;			
Best Local Similarity 83.3%; Pred. No. 2.5e+02;			
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy 1 SVDTIEY 6			
Db 772 SVDTIEY 777			

Search completed: January 10, 2003, 11:47:27
Job time : 31 secs